EFFECT OF DILUTION BOTTLE MIXING METHODS ON PLATE COUNTS OF RAW-MILK BACTERIA[1]

C. N. HUHTANEN[2], A. R. BRAZIS[3], W. L. ARLEDGE[4], E. W. COOK[5], C. B. DONNELLY[6], R. E. GINN[7], J. N. MURPHY[8], H. E. RANDOLPH[9], E. L. SING[10], AND D. I. THOMPSON[11]

(Received for publication December 17, 1969)

ABSTRACT

Forty-six raw milk samples were analyzed for plate counts at 32 C by eight different laboratories; each using three mixing procedures for the initial dilution. These were: five inversions in a 5 sec period, 15 inversions in a 15 sec period, and the "standard" method of twenty-five, 1 ft long, vertical cycles in a 7 sec period. The standard method gave the highest bacterial counts (71.1 x 10^3/ml average) the 15-15 method was second highest (60.4 x 10^3/ml average) and the 5-5 method was lowest (57.8 x 10^3/ml average). The standard method gave significantly higher ($P<0.01$) bacterial counts than the other two. The inversion methods were not significantly different from each other.

Tests of reproducibility (pooled average variances for each method) did not show any significant differences between mixing methods. There were significant differences in reproducibility between laboratories. There was evidence of interaction between mixing methods by samples and mixing methods by investigators.

Standard Methods (1) recommends a certain method of mixing dilution bottles for plating bacteria from milk. The method consists of twenty-five, 1 ft long, vertical cycles in a 7 sec period. Although the origin of the specific features of this method has been documented in the 4th edition of *Standard Methods* (1923) it differs considerably from that described in the original research article *(2)*. This method was believed to produce results of analyses which were more reproducible from laboratory to laboratory. The method is vigorous enough to be exhausting, especially for women, when large numbers of samples are being tested. The present study was undertaken to determine whether less vigorous mixing techniques would give results comparable to the "standard" method.

A blender procedure for mixing milk has been described by Wanser and Hartman *(9)*. They found an average increase in "total" plate count of 44% when either the raw milk or the initial dilution was blended for 30 sec to 1 min. This method would not be applicable to the routine study of large numbers of raw milk samples. Hartman and Huntsberger *(7)*, studied various factors influencing the microbial count of frozen foods, including the degree of mixing of dilution blanks. They found significant differences between workers and degree of shaking and also found a worker-shaking interaction. These effects were observed after a Waring blendor homogenization of the original frozen food.

The worker-to-worker reproducibility has been reported for split samples of egg salad by Messinger *(8)* and for milk by Donnelly et al. *(4, 5)*. The concept of using split samples is statistically sound since one great source of variation, between samples, is greatly reduced or eliminated. Split samples also are a great help in detecting "outliers" as shown by Donnelly et al. *(4, 5)*. The procedure used in this study was based on an analysis of variance for obtaining significant differences between mixing methods with each investigator choosing his own milk samples. This meant that the "between samples" variation would be larger than the split samples but this would be compensated for by analyzing more samples.

MATERIALS AND METHODS

Mixing techniques

Eight different laboratories participated in this study. Each secured its own raw milk samples, usually from farm bulk tanks. The assay methods were those recommended by *Standard Methods (1)* with the exception of two methods of dilution bottle mixing. Three techniques were compared for their efficacy in enumerating raw-milk bacteria: the "standard" method of twenty-five, 1 ft long vertical cycles in 7 sec; five inversions in 5 sec (5-5 method); and 15 inversions in 15 sec

---

[1]A contribution from the Subcommittee for the Examination of Milk and Milk Products, Applied Laboratory Methods Committee, International Association of Milk, Food, and Environmental Sanitarians, Inc.

[2]Eastern Utilization Research and Development Division, U. S. Dept. of Agriculture, Philadelphia, Pa. 19118 (Chairman of the Subcommittee).

[3]Food and Drug Administration, 1090 Tusculum Ave., Cincinnati, Ohio. 45226 (Chairman of the Applied Laboratory Methods Committee).

[4]Southeast Milk Sales Association, POB 1099, Bristol, Va. 24201.

[5]Quality Control Laboratory, Industrial Highway, Southampton, Pa. 18966.

[6]Food and Drug Administration, 1090 Tusculum Ave., Cincinnati, Ohio 45226.

[7]Quality Control Laboratory, 2274 Como Avenue West, St. Paul, Minn. 55108.

[8]Texas State Dept. of Health, Austin, Texas 78756.

[9]Department of Animal Science, Texas A and M University, College Station, Texas 77843.

[10]Moseley Laboratories, 3862 E. Washington St., Indianapolis, Ind. 46201.

[11]State Laboratory of Hygiene, 437 Henry Mall, Madison, Wis. 53706.

(15-15 method). These two inversion methods were selected for comparison with the standard method since they could be readily duplicated from laboratory to laboratory and were free of any significant fatigue factor when assaying large numbers of samples. The standard method is rather tiring for persons of limited physical endurance and is more likely to vary because of subjective factors such as length of stroke, duration of stroke, physical condition of the operator, etc. The time intervals of 5 and 15 sec were considered to be in an acceptable range for routine assay purposes.

Calculation of plate counts

Each method was replicated once and duplicate plates were usually poured. The total plate counts per ml raw milk were calculated from the 10^{-2} or 10^{-3} dilutions and usually the dilution showing a count between 30 and 300 colonies per plate was accepted. In some instances neither dilution fell completely in this range. In these situations the counts from a single dilution were used to avoid dilution errors. The total counts for the samples were transformed logarithmically to normalize the distributions and to ensure more nearly equal variances between the populations studied.

Statistical analyses

Generally accepted procedures for statistical evaluation were used to determine analyses variances. A missing value for one investigator was "synthesized" by a statistical formula (3):

$$X = \frac{(k-1)(t-1)}{kB + tT - G}$$

where k is the number of blocks, t is the number of treatments, B is the total of all remaining observations in the block containing the missing observation, T is the total of remaining observations in the treatment containing the missing value, and G is the total sum of observations. The use of this synthetic observation necessitated a corresponding one degree of freedom reduction in total and error degrees of freedom in the analysis of variance table (Table 1).

The analysis of variance for determining the difference between treatments was performed using the average of two duplicate petri dishes for each replicate of the method, when duplicate plates were made by the investigator. The results of Investigator No. 2 were from single plates. Two analysts, each of whom counted the colonies on one set of duplicate plates, comprise the results reported for Investigator No. 5. These counts by the two analysts were combined for the statistical analysis. The results of Investigator No. 8 also were obtained by two analysts each of whom assayed different aliquots of the same milk samples using two replicates each and duplicate petri dishes. For samples 41 to 46 inclusive the analysis of variance of treatment effects was based on the results of Investigator No. 8, Analyst a.

The data were further analyzed to determine reproducibility between methods, investigators, samples, duplicate petri dishes, and replicate milk samples. These analyses were performed by calculating single degree-of-freedom variances between paired observations, pooling these variances, and dividing by the total number of degrees of freedom to obtain a pooled average variance. These pooled average variances were then tested by the null hypothesis against each other. All tests of significance were at the 1% level.

Results and Discussion

Average counts for different mixing methods

Table 1 presents the arithmetic averages for investigators, samples, and treatments. The averages for samples were obtained by averaging over replicate milk samples and duplicate petri dishes for a

Table 1. Averages of bacterial counts obtained by three mixing methods

Milk Sample No.	Investigator No.	Mixing method: 5-5[a] (X 10^{-3})	15-15[b] (X 10^{-3})	STD[c] (X 10^{-3})
1	1	47.8	61.0	57.5
2		18.3	19.7	18.8
3		52.2	48.2	64.2
4		88.5	86.8	84.8
5		55.8	45.0	40.8
6		53.0	51.8	65.5
Average: Investigator No. 1		52.6	52.1	55.3
7	2	6.70	7.60	13.2
8		4.10	5.00	4.70
9		76.0	58.5	120.0
10		15.0	21.1[d]	68.5
11		40.5	43.0	134.0
12		6.60	5.15	6.60
13		18.4	17.2	26.7
14		20.6	22.0	52.5
Average: Investigator No. 2		23.4	22.5	53.2
15	3	161.5	153.8	167.5
16		44.0	43.8	45.2
17		55.2	56.2	55.0
18		68.5	74.2	75.2
19		69.2	68.0	105.2
20		110.2	116.0	94.0
Average: Investigator No. 3		84.4	85.3	90.4
21	4	23.6	29.3	29.6
22		86.0	91.8	92.0
23		73.8	90.2	92.8
24		18.8	31.8	32.5
25		72.5	86.5	88.0
26		35.8	51.2	49.8
Average: Investigator No. 4		51.7	63.5	64.1
27	5	44.5	50.0	86.2
28		18.5	19.6	21.5
29		92.5	86.2	83.8
30		100.5	99.2	109.2
Average: Investigator No. 5		51.2	51.1	60.2
31	6	94.8	105.8	130.5
32		33.2	30.2	56.8
33		65.8	61.8	77.2
34		49.0	48.8	48.0
35		33.2	29.8	30.5
36		53.5	60.5	55.8
37		43.2	48.2	58.8
Average: Investigator No. 6		53.2	55.0	65.4
38	7	95.5	103.5	157.5
39		5.08	5.35	6.15
40		21.8	16.5	29.5
Average: Investigator No. 7		40.8	41.8	64.4
41	8	91.0	145.0	152.5
42		215.0	267.5	242.5
43		34.0	37.5	37.5
44		32.5	37.8	43.2
45		200.0	150.0	147.5
46		59.2	63.5	74.0
Average: Investigator No. 8		105.3	110.7	116.1
Average all investigators		58.81	61.99	72.46

[a]Five inversions in 5 sec.
[b]Fifteen inversions in 15 sec.
[c]Technique of *Standard Methods*.
[d]Based on one analysis; duplicate value missing.

Line No.	Source of variation	Degrees of freedom	Sum of squares	Mean square	F ratio	Significant $P<0.01$
A	Milk Samples	45	38.524582	0.856101	71.04	Yes
B	Investigators	7	13.646441	1.949492	2.98	No
C	Samples Within Investigators	38	24.878141	0.654688	282.92	Yes
D	Treatments	2	0.6058	0.3029	25.13	Yes
E	Standard vs 5-5 and 15-15	1	0.5788	0.5788	48.03	Yes
F	5-5 vs 15-15	1	0.0266	0.0266	2.21	No
G	Treatments Times Samples	90	1.498461	0.016650	7.19	Yes
H	Investigators Times Treatments	14	0.582598	0.041614	3.45	Yes
I	Investigators Times Standard vs Others	7	0.515226	0.073604	6.11	Yes
J	Investigators Times 5-5 vs 15-15	7	0.065028	0.009290	0.78	No
K	Treatments Times Samples/Investigators	76	0.915863	0.012051	5.21	Yes
L	Error	137	0.317034	0.0023141		
	Total	274	40.945877			

[a]The F values were derived from the following ratios: A/K, B/C, C/L, D/K, E/K, F/K, G/K, H/K, I/K, K/L, J/K

TABLE 3. POOLED SINGLE-DEGREE-OF-FREEDOM VARIANCES FOR TESTING HOMOGENEITY OF VARIANCES BETWEEN MIXING METHODS AND INVESTIGATORS[a]

Investigator number	5-5 method	df	15-15 method	df	Standard method	df	Average variance
1	0.0006301	(12)	0.0038649	(12)	0.0009321	(12)	0.0018090
3	0.0069768	(12)	0.0035150	(12)	0.0021674	(12)	0.0042197
4	0.0002650	(12)	0.0002235	(12)	0.0002990	(12)	0.0002625
6	0.0038067	(14)	0.0050020	(14)	0.0034715	(14)	0.0040934
7	0.0071378	(6)	0.0046374	(6)	0.0025987	(6)	0.0047913
8a	0.0061374	(12)	0.0022679	(12)	0.0048621	(10)	0.0043966
8b	0.0099628	(12)	0.0037494	(12)	0.0099592	(12)	0.0078905
Total		(80)		(80)		(78)	
Average Variance	0.0049881		0.0033229		0.0034700		0.0038621
F Value with average variance of 15-15 method the denominator	1.47[b]				1.1[b]		

[a]The above variances were obtained by calculating variances between duplicate plates, using logarithmically transformed colony counts.
[b]Not significantly different at 1% level.

total of four observations. An exception was the result of Investigator No. 2 where the figures represented replicate milk samples but only single plates. A value missing for Investigator No. 2 (sample 10) was "synthesized" by the technique reported above.

The standard method of mixing the dilution bottles gave the highest average counts; the 15-15 method was second highest and the 5-5 method was lowest. This table shows the great variations in average counts by mixing methods depending on the sample of milk analyzed and emphasizes the necessity of assaying a large enough number of samples to get meaningful results. The variability of the plate count method is illustrated by the results in this table where 10 of the 46 samples gave the highest counts using the 15-15 method. Eight of the 46 gave highest counts using the very gentle 5-5 method. These results are further analyzed below by analysis of variance techniques.

Investigator and sample variances

A summary of the different variances which might have been expected in an experiment of this type is shown in Table 2. The very large F ratio obtained for sample variances was not unexpected since raw milk samples are known to show great sample-to-sample variations in bacterial counts. There were no significant differences between investigators at the predetermined 1% level of significance (critical F ratio was 3.15). At a lower, 5% level of significance

the encountered ratio of mean squares of 2.98 would exceed the critical F ratio of 2.26. This indicated that the number of investigators used in this experiment was adequate and nearly optimal. If the F ratio for investigators had been very small, it might have indicated that too many investigators (or laboratories) had been used and were being "wasted."

Treatment variances

The F ratios of lines D, E, and F of Table 2 showed the treatment effects to be real and were shown to result from the higher values obtained by the standard method over the other two (line E, Table 2). The difference between the 5-5 and 15-15 methods was not significant, even at the 5% level of confidence. A study of many more samples might have shown a difference although this aspect of the problem was not important enough to warrant further consideration.

Interactions

The interactions investigated, lines G-K in Table 2, indicated that there were differences in the responses of different samples to the three mixing methods. The average values for bacterial counts from Table 1 also demonstrated this type of sample versus mixing method interaction, since not all milk samples tested gave the highest counts with the standard mixing technique. This could have been a reflection of the types of bacteria present in the milk samples. Chains of some strains of *Leuconostoc citrovorum,* for instance, were found by Goel and Marth *(6)* to be shortened when subjected to the standard shaking procedure. It is also possible that certain samples of milk contained higher amounts of agglutinins, perhaps as a result of recent udder infections—these agglutinated clumps of bacteria might show greater break-up than the normal bacterial masses in milk.

The potentially more serious interactions obtained from these studies showed that there was a highly significant difference between treatments depending on the particular investigator (line H). This indicated a possible "favoring" of one method over another. A further analysis of this "favoring" indicated that it was perhaps caused by an unconscious bias for the standard method over the other two inversion techniques (line I). These interactions, however, were completely accounted for in the statistical treatment and did not negate the conclusion that the standard method was superior to the inversion methods (line E).

Reproducibility of results of three mixing methods

An important consideration in developing new methods or in comparing one method with another is the difference in variation (reproducibility) between replicates of the methods under study. In the study reported here, these variances were determined by an analysis of the pooled and averaged variances of the three methods using the plate-to-plate differences within replicates as the source of variation. The results are shown in Table 3. The statistical null hypothesis of equal mean variances for the three methods was tested using the ratios:

$$\frac{\text{5-5 method average variance}}{\text{15-15 method average variance}}$$

and

$$\frac{\text{standard method average variance}}{\text{15-15 method average variance}}$$

A higher pooled variance was obtained with the 5-5 method; however, the F ratio of 1.47 for the 5-5 method average variance/15-15 method average variance did not exceed the critical F at the 1% level of significance. The ratio was significant at the 5% level.

It appeared therefore that the least vigorous mixing method of five inversions in 5 sec was the least reproducible between laboratories and that the other two methods were about equal.

Reproducibility between investigators

The primary purpose of this study was to investigate the effect of mixing methods on mean bacterial counts and on reproducibility between methods; however the data in Table 3 also show the pooled and averaged variances obtained for each investigator. These variances were all well within the variance of log plate counts suggested by Donnelly et al. *(4, 5),* of 0.012. The pooled variances of Table 3 show that there were great differences in precision between investigators (or laboratories, since the investigators, except 8a and 8b, were also in separate laboratories). The lowest average variance, 0.0002625, was attained by Investigator No. 4 while the highest, 0.0078905, was that of Investigator No. 8b. Investigator 8a, in the same laboratory as 8b, had a lower variance. Bartlett's and Cochran's tests for homogeneity of variances showed the between-investigator average variances to be significantly different. Inspection of the investigator average variances showed most of this difference to be due to the very low average variance of Investigator No. 4. The explanation for these variations in precision is not known but would be important, since it would be in the interest of all laboratories to adopt the procedures which would give the most reproducible results.

The wide range of precision between investigators does not necessarily negate the F test of the analysis of variance since the populations of interest were methods rather than investigators and because the F test is powerful enough to yield satisfactory results even with such widely differing population variances.

Acknowledgments

The aid of Drs. S. M. Free and W. J. Westlake, Smith, Kline and French Laboratories, Philadelphia, Pennsylvania, for statistical consultation is gratefully acknowledged.

References

1. American Public Health Association. 1967. Standard methods for the examination of dairy products, 12th ed. American Public Health Association, Inc., New York, New York.

2. Conn, H. W. 1915. Standards for determining the purity of milk. The limit of error in bacteriological milk analyses. USPHS Public Health Rep. 30:2349-2395.

3. Cox, D. R. 1958. Planning of experiments. John Wiley and Sons, Inc., New York, New York. p. 31.

4. Donnelly, C. B., E. K. Harris, L. A. Black, and K. H. Lewis. 1960. Statistical analysis of standard plate counts of milk samples split with state laboratories. J. Milk Food Technol. 23:315-319.

5. Donnelly, C. B., J. T. Peeler, and L. A. Black. 1966. Evaluation of state central milk laboratories by statistical analyses of standard plate counts. J. Milk Food Technol. 29:19-24.

6. Goel, M. C., and E. H. Marth. 1969. Chain length of *Leuconostoc citrovorum* modified by dilution and shaking procedures of the plate count. J. Dairy Sci. 52:1941-1947.

7. Hartman, P. A., and D. V. Huntsberger. 1961. Influence of subtle differences in plating procedure on bacterial counts of prepared frozen foods. Appl. Microbiol. 9:32-38.

8. Messinger, H. B. 1963. Statistical analysis of standard plate counts of a food sample split among laboratories. J. Milk Food Technol. 26:328-331.

9. Wanser, B. E., and P. A. Hartsman. 1962. Agitation of milk samples for colony counts of raw milk. J. Milk Food Technol. 25:5-6.